

shown in SEQ ID NO:2. In addition, the nucleotide sequence of the coding region is depicted in SEQ ID NO:3.

Accordingly, in one aspect, the invention features a nucleic acid molecule which encodes a 32544 protein or polypeptide, e.g., a biologically active portion of the 32544 protein. In a preferred embodiment, the isolated nucleic acid molecule encodes a polypeptide having the amino acid sequence of SEQ ID NO:2. In other embodiments, the invention provides an isolated 32544 nucleic acid molecule having the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number \_\_\_\_\_. In still other embodiments, the invention provides nucleic acid molecules that are substantially identical (e.g., naturally occurring allelic variants) to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number \_\_\_\_\_. In other embodiments, the invention provides a nucleic acid molecule which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number \_\_\_\_\_, wherein the nucleic acid encodes a full length 32544 protein or an active fragment thereof.

In a related aspect, the invention further provides nucleic acid constructs which include a 32544 nucleic acid molecule described herein. In certain embodiments, the nucleic acid molecules of the invention are operatively linked to native or heterologous regulatory sequences. Also included, are vectors and host cells containing the 32544 nucleic acid molecules of the invention e.g., vectors and host cells suitable for producing 32544 nucleic acid molecules and polypeptides.

In another related aspect, the invention provides nucleic acid fragments suitable as primers or hybridization probes for the detection of 32544-encoding nucleic acids.

In still another related aspect, isolated nucleic acid molecules that are antisense to a 32544 encoding nucleic acid molecule are provided.

In another aspect, the invention features, 32544 polypeptides, and biologically active or antigenic fragments thereof that are useful, e.g., as reagents or targets in assays applicable to treatment and diagnosis of 32544-mediated or related disorders. In another embodiment, the invention provides 32544 polypeptides having a 32544 activity. Preferred polypeptides

are 32544 proteins including at least one domain, e.g., a PLC-X domain (from about amino acids 323-468 of SEQ ID NO:2), a PLC-Y domain (from about amino acids 621-736 of SEQ ID NO:2), a calcium binding (C2) domain (from about amino acids 756-848 of SEQ ID NO:2), or a pleckstrin homology (PH) domain (from about amino acids 44-151 of SEQ ID NO:2), and, preferably, having a 32544 activity, e.g., an activity as described herein, e.g., the ability to catalyze the hydrolysis of phosphatidyl-inositol-4,5-bisphosphate (PIP2) producing diacylglycerol and inositol 1,4,5-trisphosphate (IP3).

In other embodiments, the invention provides 32544 polypeptides, e.g., a 32544 polypeptide having the amino acid sequence shown in SEQ ID NO:2; the amino acid sequence encoded by the cDNA insert of the plasmid deposited with ATCC Accession Number \_\_\_\_; an amino acid sequence that is substantially identical to the amino acid sequence shown in SEQ ID NO:2; or an amino acid sequence encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the sequence of the DNA insert of the plasmid deposited with ATCC Accession Number \_\_\_\_, wherein the nucleic acid encodes a full length 32544 protein or an active fragment thereof.

In a related aspect, the invention further provides nucleic acid constructs which include a 32544 nucleic acid molecule described herein.

In a related aspect, the invention provides 32544 polypeptides or fragments operatively linked to non-32544 polypeptides to form fusion proteins.

In another aspect, the invention features antibodies and antigen-binding fragments thereof, that react with, or more preferably specifically bind 32544 polypeptides.

In another aspect, the invention provides methods of screening for compounds that modulate the expression or activity of the 32544 polypeptides or nucleic acids.

In still another aspect, the invention provides a process for modulating 32544 polypeptide or nucleic acid expression or activity, e.g. using the screened compounds. In certain embodiments, the methods involve treatment of conditions related to aberrant activity or expression of the 32544 polypeptides or nucleic acids, such as conditions involving aberrant or deficient cellular proliferation or differentiation.

09927112.081001

A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of 32544 (e.g., the sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_) without abolishing or more preferably, without substantially altering a biological activity, whereas an "essential" amino acid residue results in such a change. For example, amino acid residues that are conserved among the polypeptides of the present invention, e.g., those present in the phospholipase family domain, are predicted to be particularly unamenable to alteration.

A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in a 32544 protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of a 32544 coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for 32544 biological activity to identify mutants that retain activity. Following mutagenesis of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

As used herein, a "biologically active portion" of a 32544 protein includes a fragment of a 32544 protein which participates in an interaction between a 32544 molecule and a non-32544 molecule. Biologically active portions of a 32544 protein include peptides comprising amino acid sequences sufficiently homologous to or derived from the amino acid sequence of the 32544 protein, e.g., the amino acid sequence shown in SEQ ID NO:2, which include less amino acids than the full length 32544 proteins, and exhibit at least one activity of a 32544

fragments suitable for use as primers, e.g., PCR primers for the amplification or mutation of nucleic acid molecules.

In one embodiment, an isolated nucleic acid molecule of the invention includes the nucleotide sequence shown in SEQ ID NO:1, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, or a portion of any of these nucleotide sequences. In one embodiment, the nucleic acid molecule includes sequences encoding the human 32544 protein (i.e., "the coding region", from nucleotides 435-4055 of SEQ ID NO:1, not including the terminal codon), as well as 5' untranslated sequences (nucleotides 1-434 of SEQ ID NO:1). Alternatively, the nucleic acid molecule can include only the coding region of SEQ ID NO:1 (e.g., nucleotides 435-4055 of SEQ ID NO:1, corresponding to SEQ ID NO:3) and, e.g., no flanking sequences which normally accompany the subject sequence. In another embodiment, the nucleic acid molecule encodes a sequence corresponding to the mature protein of SEQ ID NO:2.

In another embodiment, an isolated nucleic acid molecule of the invention includes a nucleic acid molecule which is a complement of the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, or a portion of any of these nucleotide sequences. In other embodiments, the nucleic acid molecule of the invention is sufficiently complementary to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_ such that it can hybridize to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, thereby forming a stable duplex.

In one embodiment, an isolated nucleic acid molecule of the present invention includes a nucleotide sequence which is at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more homologous to the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_. In the case of an isolated nucleic acid molecule which is longer than or equivalent in length to the reference sequence, e.g., SEQ ID NO:1, or SEQ ID NO:3, the comparison is made with the full length of the reference sequence. Where the isolated nucleic acid molecule is shorter

than the reference sequence, e.g., shorter than SEQ ID NO:1, or SEQ ID NO:3, the comparison is made to a segment of the reference sequence of the same length (excluding any loop required by the homology calculation).

5. 32544 Nucleic Acid Fragments

A nucleic acid molecule of the invention can include only a portion of the nucleic acid sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_. For example, such a nucleic acid molecule can include a fragment which can be used as a probe or primer or a fragment encoding a portion of a 32544 protein, e.g., an immunogenic or biologically active portion of a 32544 protein. A fragment can comprise e.g., nucleotides 275 to 598 of SEQ ID NO:1, which encodes a PH domain of human 32544; nucleotides 1112 to 1549 of SEQ ID NO:1, which encodes the PLC-X domain of human 32544; or nucleotides 2006 to 2353 of SEQ ID NO:1, which encodes the PLC-Y domain of human 32544. The nucleotide sequence determined from the cloning of the 32544 gene allows for the generation of probes and primers designed for use in identifying and/or cloning other 32544 family members, or fragments thereof, as well as 32544 homologues, or fragments thereof, from other species.

In another embodiment, a nucleic acid includes a nucleotide sequence that includes part, or all, of the coding region and extends into either (or both) the 5' or 3' noncoding region. Other embodiments include a fragment which includes a nucleotide sequence encoding an amino acid fragment described herein. Nucleic acid fragments can encode a specific domain or site described herein or fragments thereof, particularly fragments thereof which are at least 150 amino acids in length. Fragments also include nucleic acid sequences corresponding to specific amino acid sequences described above or fragments thereof. Nucleic acid fragments should not to be construed as encompassing those fragments that may have been disclosed prior to the invention.

A nucleic acid fragment can include a sequence corresponding to a domain, region, or functional site described herein. A nucleic acid fragment can also include one or more domain, region, or functional site described herein. Thus, for example, the nucleic acid fragment can include a PH domain and a C2 domain. In a preferred embodiment the fragment is at least, 50, 100, 200, 300, 400, 500, 600, 700, or 900 base pairs in length.

32544 probes and primers are provided. Typically a probe/primer is an isolated or purified oligonucleotide. The oligonucleotide typically includes a region of nucleotide sequence that hybridizes under stringent conditions to at least about 7, 12 or 15, preferably about 20 or 25, more preferably about 30, 35, 40, 45, 50, 55, 60, 65, or 75 consecutive nucleotides of a sense or antisense sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, or of a naturally occurring allelic variant or mutant of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_.

In a preferred embodiment the nucleic acid is a probe which is at least 5 or 10, and less than 200, more preferably less than 100, or less than 50, base pairs in length. It should be identical, or differ by 1, or less than in 5 or 10 bases, from a sequence disclosed herein. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

A probe or primer can be derived from the sense or anti-sense strand of a nucleic acid which encodes: PH domain (from about amino acids 44-151 of SEQ ID NO:2); a PLC-X domain (from about amino acids 323-468 of SEQ ID NO:2); a PLC-Y domain (from about amino acids 621-736 of SEQ ID NO:2); or a C2 domain (from about amino acids 756-848 of SEQ ID NO:2). In another embodiment a set of primers is provided, e.g., primers suitable for use in a PCR, which can be used to amplify a selected region of a 32544 sequence. The primers should be at least 5, 10, or 50 base pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differs by one base from a sequence disclosed herein or from a naturally occurring variant. E.g., primers suitable for amplifying a domain or region described herein are provided, e.g., all or a portion of any of the following regions: a PLC-X domain (from about amino acids 323-468 of SEQ ID NO:2), a PLC-Y (from about amino acids 621-736 of SEQ ID NO:2), a C2 domain (from about amino acids 756-848 of SEQ ID NO:2), or a PH domain (from about amino acids 44-151 of SEQ ID NO:2).

In another embodiment a set of primers is provided, e.g., primers suitable for use in a PCR, which can be used to amplify a selected region of a 32544 sequence, e.g., a region

described herein. The primers should be at least 5, 10, or 50 base pairs in length and less than 100, or less than 200, base pairs in length. The primers should be identical, or differs by one base from a sequence disclosed herein or from a naturally occurring variant. E.g., primers suitable for amplifying all or a portion of any of the following regions are provided:

5 a phospholipase family domain (e.g., about nucleotides 1401-1838 of SEQ ID NO:1).

A nucleic acid fragment can encode an epitope bearing region of a polypeptide described herein.

A nucleic acid fragment encoding a "biologically active portion of a 32544 polypeptide" can be prepared by isolating a portion of the nucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, which encodes a polypeptide having a 32544 biological activity (e.g., the biological activities of the 32544 proteins as described herein), expressing the encoded portion of the 32544 protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the 32544 protein. For example, a nucleic acid fragment encoding a biologically active portion of 32544 includes a phospholipase family domain (e.g., about nucleotides 1401-1838 of SEQ ID NO:1). A nucleic acid fragment encoding a biologically active portion of a 32544 polypeptide, may comprise a nucleotide sequence which is greater than 300-1200 or more nucleotides in length.

In preferred embodiments, nucleic acids include a nucleotide sequence which is about 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200 nucleotides in length and hybridizes under stringent hybridization conditions to a nucleic acid molecule of SEQ ID NO:1, or SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_.

#### 32544 Nucleic Acid Variants

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence shown in SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_. Such differences can be due to degeneracy of the genetic code (and result in a nucleic acid which encodes the same 32544 proteins as those encoded by the nucleotide sequence disclosed

herein. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence which differs, by at least 1, but less than 5, 10, 20, 50, or 100 amino acid residues that shown in SEQ ID NO:2. If alignment is needed for this comparison the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

Nucleic acids of the inventor can be chosen for having codons, which are preferred, or non preferred, for a particular expression system. E.g., the nucleic acid can be one in which at least one codon, at preferably at least 10%, or 20% of the codons has been altered such that the sequence is optimized for expression in E. coli, yeast, human, insect, or CHO cells.

Nucleic acid variants can be naturally occurring, such as allelic variants (same locus), homologs (different locus), and orthologs (different organism) or can be non-naturally occurring. Non-naturally occurring variants can be made by mutagenesis techniques, including those applied to polynucleotides, cells, or organisms. The variants can contain nucleotide substitutions, deletions, inversions and insertions. Variation can occur in either or both the coding and non-coding regions. The variations can produce both conservative and non-conservative amino acid substitutions (as compared in the encoded product).

In a preferred embodiment, the nucleic acid differs from that of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_, e.g., as follows: by at least one but less than 10, 20, 30, or 40 nucleotides; at least one but less than 1%, 5%, 10% or 20% of the in the subject nucleic acid. If necessary for this analysis the sequences should be aligned for maximum homology. "Looped" out sequences from deletions or insertions, or mismatches, are considered differences.

Orthologs, homologs, and allelic variants can be identified using methods known in the art. These variants comprise a nucleotide sequence encoding a polypeptide that is 50%, at least about 55%, typically at least about 70-75%, more typically at least about 80-85%, and most typically at least about 90-95% or more identical to the amino acid sequence shown in SEQ ID NO:2 or a fragment of this sequence. Such nucleic acid molecules can readily be obtained as being able to hybridize under stringent conditions, to the nucleotide sequence shown in SEQ ID NO:3 or a fragment of this sequence. Nucleic acid molecules corresponding to orthologs, homologs, and allelic variants of the 32544 cDNAs of the invention can further be isolated by

mapping to the same chromosome or locus as the 32544 gene. Preferred variants include those that are correlated with phospholipase family activity.

Allelic variants of 32544, e.g., human 32544, include both functional and non-functional proteins. Functional allelic variants are naturally occurring amino acid sequence variants of the 32544 protein within a population that maintain the ability to function similar to the 32544 protein, e.g., to hydrolyze phosphatidylinositol 4,5-bisphosphate to produce inositol 1,4,5-trisphosphate and diacylglycerol.

Functional allelic variants will typically contain only conservative substitution of one or more amino acids of SEQ ID NO:2, or substitution, deletion or insertion of non-critical residues in non-critical regions of the protein. Non-functional allelic variants are naturally-occurring amino acid sequence variants of the 32544, e.g., human 32544, protein within a population that do not have the ability to hydrolyze phosphatidylinositol. Non-functional allelic variants will typically contain a non-conservative substitution, a deletion, or insertion, or premature truncation of the amino acid sequence of SEQ ID NO:2, or a substitution, insertion, or deletion in critical residues or critical regions of the protein.

Moreover, nucleic acid molecules encoding other 32544 family members and, thus, which have a nucleotide sequence which differs from the 32544 sequences of SEQ ID NO:1, SEQ ID NO:3, or the nucleotide sequence of the DNA insert of the plasmid deposited with ATCC as Accession Number \_\_\_\_\_ are intended to be within the scope of the invention.

#### Antisense Nucleic Acid Molecules, Ribozymes and Modified 32544 Nucleic Acid Molecules

In another aspect, the invention features, an isolated nucleic acid molecule which is antisense to 32544. An "antisense" nucleic acid can include a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. The antisense nucleic acid can be complementary to an entire 32544 coding strand, or to only a portion thereof (e.g., the coding region of human 32544 corresponding to SEQ ID NO:3). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding 32544 (e.g., the 5' and 3' untranslated regions).